Econophysics - written exam 23 January 2015 with solutions

Docent: Dr. Diego Garlaschelli

You should return these pages completed. Please write your answers to all questions by filling the boxes. Motivate each of your answers. Attach all the additional sheets required to prove your results. Answers are ignored if there is no explanation or proof. Please write clearly and in English. Please don't forget to write your name and student ID below. Good luck!

NAME: . . . .

SURNAME: . . . .

STUDENT ID: . . . .

PROBLEM 1 (3 POINTS)

Consider a random process generating the time series of a single financial asset. The time series is recorded for T discrete time steps, each of equal duration Δt (e.g. $\Delta t = 1$ sec, 1 day, etc.). The random variable denoting the increment (log-return) of the time series at the i-th time step is denoted as X_i, so that a realization of the entire time series is a realization of the process $(X_1, X_2, \ldots, X_T)$. The only thing we assume about the T random variables $\{X_i\}_{i=1}^T$ is that they have the same expected value $\langle X_i \rangle = \mu \; \forall i$. Let us denote the expected product of X_i and X_j as $\langle X_i X_j \rangle$ and define the $T \times T$ matrix $\mathbf{M}$ as the matrix with entries $M_{ij} = \langle X_i X_j \rangle$. Now assume that $\mathbf{M}$ has the form

$$\mathbf{M} = \begin{pmatrix} a & b & \mu^2 & \mu^2 & \mu^2 & \mu^2 & \cdots & \mu^2 \\ b & a & b & \mu^2 & \mu^2 & \mu^2 & \cdots & \mu^2 \\ \mu^2 & b & a & b & \mu^2 & \mu^2 & \cdots & \mu^2 \\ \mu^2 & \mu^2 & b & a & b & \mu^2 & \cdots & \mu^2 \\ \vdots & & & & & & & \vdots \\ \mu^2 & \cdots & \mu^2 & b & a & b & \mu^2 & \mu^2 \\ \mu^2 & \cdots & \mu^2 & \mu^2 & b & a & b & \mu^2 \\ \mu^2 & \cdots & \mu^2 & \mu^2 & \mu^2 & b & a & b \\ \mu^2 & \cdots & \mu^2 & \mu^2 & \mu^2 & \mu^2 & b & a \end{pmatrix} \tag{1}$$

1. Determine the variance σ^2 of X_i. Determine the condition on a ensuring that $\sigma^2 \geq 0$.

$\sigma^2 = a - \mu^2$ (since $\sigma^2 = \langle X_i^2 \rangle - \langle X_i \rangle^2 = M_{ii} - \mu^2$)

Condition on a: $a \geq \mu^2$ (follows from the requirement $\sigma^2 \geq 0$)

2. Determine the covariance of two <u>non-consecutive</u> increments X_i and X_{i+t} with $t > 1$.

$\mathrm{Cov}[X_i, X_{i+t}] = \langle X_i X_{i+t} \rangle - \langle X_i \rangle \langle X_{i+t} \rangle = M_{i,i+t} - \mu^2 = \mu^2 - \mu^2 = 0$ $(t > 1)$

3. Determine the covariance of two <u>consecutive</u> increments X_i and X_{i+1}.

$\mathrm{Cov}[X_i, X_{i+1}] = \langle X_i X_{i+1} \rangle - \langle X_i \rangle \langle X_{i+1} \rangle = M_{i,i+1} - \mu^2 = b - \mu^2$

4. Determine the correlation coefficient $\rho \equiv \mathrm{Corr}[X_i, X_{i+1}]$ of two consecutive increments X_i and X_{i+1}. Given the condition on a found above in point 1, find the condition on b ensuring that $\rho \in [-1, +1]$.

$$\rho = \mathrm{Corr}[X_i, X_{i+1}] = \frac{\mathrm{Cov}[X_i, X_{i+1}]}{\sqrt{\mathrm{Var}[X_i]\mathrm{Var}[X_{i+1}]}} = \frac{b - \mu^2}{\sigma^2} = \frac{b - \mu^2}{a - \mu^2}$$

Condition on b: $2\mu^2 - a \leq b \leq a$ ($\rho \leq +1$ implies $b \leq a$; $\rho \geq -1$ implies $b \geq 2\mu^2 - a$)

5. Can you find condition(s) on a and/or b that ensure that all the T increments are mutually uncorrelated? Why?

Yes, it suffices to require that $\mathrm{Corr}[X_i, X_j]$=0 $\forall i \neq j$. This implies $\rho = 0$, i.e. $b = \mu^2$.

6. Can you find condition(s) on a and/or b that ensure that all the T increments are mutually independent? Why?

No, because even when all correlations are zero, the random variables can be dependent (independence implies no correlation, but not viceversa).

PROBLEM 2 (3 POINTS)

With reference to Problem 1, now consider the random variable

$$Y \equiv \sum_{i=1}^{T} X_i \tag{2}$$

describing the total increment of the time series after T time steps. The matrix $\mathbf{M}$ is the same as in eq.(1), with generic parameters a, b, μ^2 (subject to the conditions discussed previously).

1. Calculate the variance of Y as a function of a, b, μ^2, T.

$$\mathrm{Var}[Y] = \sum_{i=1}^{T}\sum_{j=1}^{T} \mathrm{Cov}[X_i, X_j] = \sum_{i=1}^{T}\sum_{j=1}^{T}(M_{ij} - \mu^2) = T(a-\mu^2) + 2(T-1)(b-\mu^2)$$

2. Now consider the case where the T increments are mutually uncorrelated, and let Y_0 denote Y in this particular case. Using your answer in point 5 of Problem 1, calculate the variance of Y_0.

$\mathrm{Var}[Y_0] = T(a-\mu^2)$ (since $b = \mu^2$ when increments are uncorrelated)

3. Express the ratio $\mathrm{Var}[Y]/\mathrm{Var}[Y_0]$ as a function of only T and ρ (where ρ is the 'consecutive' correlation coefficient determined previously in point 4 of Problem 1) and calculate its value in the limit of infinitely long time series.

$$r \equiv \lim_{T\to+\infty} \frac{\mathrm{Var}[Y]}{\mathrm{Var}[Y_0]} = \lim_{T\to+\infty} \frac{T(a-\mu^2) + 2(T-1)(b-\mu^2)}{T(a-\mu^2)} = \lim_{T\to+\infty}\left[1 + 2\rho\,\frac{T-1}{T}\right] = 1 + 2\rho$$

4. Calculate the value of the consecutive correlation ρ required to ensure, in the limit of infinite T, that the total variance of the correlated time series is twice as big as that of the uncorrelated one, i.e. $r = 2$.

$\rho = \frac{r-1}{2} = \frac{1}{2}$ when $r = 2$

5. Calculate the value of the consecutive correlation ρ required to ensure, in the limit of infinite T, that the total variance of the correlated time series is negligible with respect to that of the uncorrelated one, i.e. $r = 0$.

$\rho = \frac{r-1}{2} = -\frac{1}{2}$ when $r = 0$

6. Discuss the (necessary) conditions on a and μ required to find the distribution of Y via the Central Limit Theorem.

$\mu < +\infty$ (to apply the CLT to a sum of variables, all means must be finite)
$a < +\infty$ (to apply the CLT to a sum of variables, all variances must be finite)

7. Discuss the (sufficient) condition on b required to find the distribution of Y via the Central Limit Theorem.

$b = \mu^2$ (if the CLT can be applied, it means that the variables in the sum are all independent)

8. Define a critical timescale Δt_c such that, for $\Delta t \ll \Delta t_c$, the model defined in Problem 1 is a bad model for the autocorrelation of financial log-returns in most empirical time series, while for $\Delta t \gg \Delta t_c$ the model is approximately good. Determine the order of magnitude of Δt_c and explain why.

$\Delta t_c \approx 1min$, because real log-returns are autocorrelated over shorter timescales and uncorrelated over longer timescales

PROBLEM 3 (3 POINTS)

Now assume that the matrix $\mathbf{M}$ of eq.(1) describes the probability of connection in an undirected network, i.e. M_{ij} is the probability that nodes i and j are connected, and T is the number of nodes. Ignore all the conditions on a, b, μ considered previously.

1. Assume $a = 0$, $b = 1$ and $\mu = 0$. What does the resulting network model look like?

   Network model: unidimensional chain of T nodes

2. Assume $a = 0$, $0 \leq b \leq 1$, $\mu = \sqrt{b}$. What does the resulting network model look like? What is the interpretation of b in this case? Describe what happens as b varies from 0 to 1.

   Network model: Erdő-Rényi random graph

   Interpretation of b: connection probability in the Erdő-Rényi model

   As b varies from 0 to 1, the random graph changes from an empty graph to a full graph (optional: undergoing a percolation transition at the critical value $b_c = 1/T$)

3. Assume $a = 0$, $b = 1 - p$. Calculate the value of μ that ensures that, irrespective of the value of p, the expected number $\langle L \rangle$ of links in this model coincides with the expected number of links in the model in point 1 above. What does the resulting network model look like? What is the interpretation of p? Describe qualitatively what happens to the expected clustering coefficient and average distance as p varies from 0 to 1.

   $\mu = \sqrt{\dfrac{2p}{T-2}}$: follows from equating $\langle L \rangle = \sum_{i<j} M_{ij} = (T-1)\left[1 - p + \mu^2 \left(\dfrac{T}{2} - 1\right)\right]$ and $T - 1$

   Network model: Watts-Strogatz "small-world" model (optional: apart from the openness of the chain)

   Interpretation of p: rewiring probability

   As p varies from 0 to 1, the network changes from a unidimensional lattice to a random graph (optional: with an intermediate "small-world" region).

4. Indicate which (if any) of the above network model(s) can generate an average vertex-vertex distance consistent with that of most empirical economic networks, and which model(s) can do the same for the degree distribution.

   Good model(s) for distance: 2 and 3 Good model(s) for degree distribution: none

5. Indicate which network model(s), among the three ones defined above, can generate a power-law distribution of log-returns if used as the agent-agent interaction graph in the Cont-Bouchaud model. Indicate the specific value of the parameters a, b, μ for which this is true.

   Model: Erdős-Rényi random graph (2)

   $a = 0$ $b = T^{-1}$ (only at the percolation threshold) $\mu = \sqrt{b} = T^{-1/2}$